知识管理论坛　ISSN 2095-5472　CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

基于科研过程的科学数据安全行为研究

◎杨燕[1]　阮建海[2]
1. 西南大学计算机与信息科学学院　重庆 400715
2. 西南大学图书馆　重庆 400715

摘要：[目的 / 意义] 通过界定科学数据安全行为的概念，分析科研活动不同阶段应如何保证科学数据安全，对规范科研人员的科学数据安全行为、促进科学数据安全管理、制定科学数据安全政策、完善科学数据安全管理体系具有重要意义。[方法 / 过程] 采用网络调研法，整合信息安全行为和科学数据安全管理的定义，提出科学数据安全行为的概念。在此基础上，从理论出发，结合科学数据管理与共享的实践经验，系统阐述如何通过个人行为保障科学数据安全。[结果 / 结论] 现阶段科学数据安全管理主要针对科学数据管理平台，从行为学视角探讨科学数据安全的研究尚处于起步阶段。本文基于科研过程，从两方面对如何规范科研人员的行为以提升科学数据安全进行分析：其一，对存在科学数据安全风险的阶段，结合科学数据管理的政策对此提出针对性的改进策略；其二，对不存在明显安全风险的阶段，根据国内外科学数据管理的实践经验，提出建议。

关键词：科学数据安全行为；科学数据；信息安全行为；科学数据管理
分类号：G251

引用格式：杨燕 , 阮建海 . 基于科研过程的科学数据安全行为研究 [J/OL]. 知识管理论坛 , 2019, 4(4): 218-231[引用日期]. http://www.kmf.ac.cn/p/181/.

科学数据是科研观测、科学研究活动的成果，是科技创新和国家发展的重要资源[1-2]。鉴于科学数据共享对科技创新能力、科技整体水平以及社会经济发展有促进作用[3]，全球范围内掀起了科学数据开放获取的热潮，不少国际组织、国家、机构建立起了数据管理中心，如世界数据中心[4]、澳大利亚国家数据服务中心[5]、英国数字管理中心[6]以及以国外高校图书馆（如爱丁堡大学图书馆[7]、普林斯顿大学图书馆[8]等）为首建立的数据管理中心。科学数据开放共享在促进科学技术快速发展的同时，也造成了“数据获取”和“数据安全及隐私保护”之间的不平衡，以致出现了一系列的科学数据安全问题。科学数据安全不仅关乎个人信息安全，同时还关乎国家安全[9]。为了保护科学数据安全，各国从宏观层面制定了相关政策，比如英国 1998 年颁布的数据保护法案（Data Protection Act）[10]，欧盟委员会（EC）发布的《EC

基金项目：本文系重庆市科学规划项目“科学数据安全行为研究”（项目编号 : 2017YBCB068）研究成果之一。
作者简介：杨燕（ORCID: 0000-0001-5152-8348），硕士研究生，E-mail: 1293553756@qq.com；阮建海（ORCID: 0000-0001-6171-7441），研究馆员，教授，博士，博士生导师。
收稿日期：2019-02-25　　发表日期：2019-08-01　　本文责任编辑：刘远颖

知识管理论坛，2019(4):218-231
DOI：10.13266/j.issn.2095-5472.2019.024

对访问与保存科学信息的建议》[11]，美国颁布的《健康保险携带和责任法案》（Health Insurance Portability and Accountability Act，HIPAA）[12]。在微观层面，相关科研项目基金会要求他们所资助的研究者提交可以为科研成果提供佐证的科研数据，同时对数据的存储和保存以及科学数据的处理方式都做了相应规定[13]。

根据2018全球信息安全调查显示，绝大多数安全事故均是由人的不安全行为引发的[14]。联合信息系统委员会（Joint Information Systems Committee）指出潜在的科学数据安全风险通常不是由所采取的技术造成的，而是由于研究人员不恰当的行为导致[15]。例如，2018年引起全球关注的Facebook数据泄露事件，其缘由是剑桥大学心理测量中心教授将从Facebook收集的用于性格分析的用户数据泄露给了剑桥分析公司[16]。目前，对科学数据安全的探讨主要集中在政策和技术层面，而科研人员对政策的遵循和技术的采纳程度还受到个人主观意愿的影响。所以学者开始关注从行为学的视角，探究如何让科研人员能主动、有意识地保护科学数据安全。

鉴于此，为了规范科研人员的科学数据安全管理方式，让科学数据更好地为国家科技创新、经济社会发展和国家安全提供支撑[17]，笔者从现实需求出发，提出科学数据安全行为这一概念。在科研过程中，科学数据安全主要受到科研人员行为的影响。在科学数据生命周期中，科研人员同时扮演着数据的生产者、使用者、管理者和监管者等多重角色[18]。科研人员的多重身份与科学数据形式的动态变化特征为科学数据安全管理增加了难度，规范科研人员的行为有利于实现科学数据的有效管理。所以笔者基于科研过程阐述不同阶段存在哪些科学数据安全风险，以及应该从哪些方面可以规范科研人员的科学数据安全行为，进一步丰富科学数据安全管理的研究内容，促进科学数据管理与共享的发展，并为科学数据安全管理提供指导和建议，为科学数据管理的政策制定奠定理论基础。

1 科学数据安全行为概念的界定

科学数据安全行为是本文的研究对象，从已有的研究成果来看，科学数据安全行为方面的研究还相对较少，针对科学数据安全行为的概念界定尚未形成统一的认识。从科学数据安全行为产生、发展的脉络看，科学数据安全行为主要涉及信息安全行为和科学数据安全管理两个主题领域。就科学数据安全行为的体系构成而言，科学数据安全行为是信息安全行为与科学数据安全管理二者的组合体，同时隶属于信息安全行为与科学数据安全管理的研究范畴。从科学数据安全行为的理论渊源看，科学数据安全行为的概念是由信息安全行为和科学数据安全管理两个概念有机整合而引申出来的一个复合型概念。因此，明晰信息安全行为与科学数据安全管理两个基本概念是提出科学数据安全行为的基础和前提。

1.1 信息安全行为的定义

有关人为因素对计算机运行可靠性的影响的相关研究可以追溯到20世纪50年代[19]，直到20世纪90年代系统科学的兴起，对信息安全行为的研究才开始引起学者的广泛关注[20]。以C. Wood等为代表的学者最早指出人为因素对信息系统安全具有潜在影响[21]。90年代中后期，互联网技术的快速发展，使得维系信息系统运转的软件、硬件变得更加复杂，学者开始致力于探讨如何充分利用技术的有用性减少人为差错带来的安全风险[22]。21世纪初，随着信息系统在各个领域的应用，研究发现技术的应用并不能完全保证信息的安全[23]，人为因素是导致信息安全问题的最根本的因素[24]，对信息安全的研究重心开始向个体行为转移。从国内外的研究成果看，信息安全行为的研究已经形成了相对完善的研究体系。关于信息安全的定义有如下解释：J. M. Stanton等从个人行为角度出发认为信息安全行为是指维护信息系统有效性、机密性和完整性的人的行为集合[25]；S.



Dzazali 等从组织角度出发认为信息安全行为是企业员工为避免组织财产遭遇损失及机密信息被曝光的行为 [26]；H. Liang 等认为信息威胁规避行为，是指当用户感知到信息存在安全威胁时采取有效的防护措施的行为 [27]；J. Wirtz 等提出网络隐私保护行为，指的是用户通过技术手段，确定网络安全性来保护自己的信息及隐私免受侵犯的行为 [28]。可以看出，信息安全行为主要从行为学的视角出发，探讨如何调动个体的主观能动性，以消除信息安全威胁、保障信息安全为目的，而采取的相关有效措施的行为集合。

1.2 科学数据安全管理的含义

随着数据密集型科学研究范式的兴起，数据逐渐成为科学研究的核心 [29]。其共享和再利用被认为是科技创新和知识发现的重要驱动因素之一。随着国内外科学共享工程的开展，如何有效地对科学数据进行管理和完善，成为高校及科研机构迫切关心的问题 [30]。科学数据的综合管理是一个有机的复杂系统，包括技术基础设施的建设、社会基础设施的建立、政策框架、商务计划、人员布局等 [31]。通过相关理论研究和实践探索，学者提出数据监管，即为确保数据当前使用目的，并能用于未来再发现及再利用，从数据产生开始对其进行管理和完善的活动 [32-33]。在科学数据监管体系中，科学数据安全管理是其中的关键环节。数据安全管理是对科学数据管理中存在的安全问题进行管理的过程 [34]。科学数据安全管理的主要管理活动包括科学数据安全管理的计划、安全隐患识别、安全问题的定性定量评估、安全威胁的应对措施以及安全威胁控制等 [34]。学者通过调研国内外的科学数据安全政策发现，科学数据安全管理具体包括数据的分类规则、数据共享的限制（对数据保密级别、共享方式进行说明）、数据处理、对涉及人类参与者隐私安全保护等几个方面 [35-37]。在技术层面，现有的科学数据管理平台主要采用分布式系统架构，也有学者提出运用区块链技术实现交互数据描述、安全管理等功能 [38]。总体来说，科学数据安全问题已经引起全社会的广泛关注，目前针对科学数据安全管理的探讨主要集中在政策和技术层面。

1.3 科学数据安全行为的定义

虽然相关政策和技术都在逐步完善，但是在缺乏行为规范的环境下，仍然无法完全保证科学数据的安全。在科学数据生命周期中，科学数据在共享之前由科研人员直接管理为主，机构或科学数据管理平台协助管理为辅。所以在这个过程中科学数据的安全与科研人员的行为密不可分。信息安全行为的相关研究证明，个体行为是造成信息安全的根本因素。与信息有所区别的是，科学数据是指在科技活动（实验、观察、探测、调查等）或通过其他方式所获取的反映客观世界的本质、特征、变化规律等的原始基本数据，以及根据不同科技活动需要，进行系统加工整理的各类数据集 [39]。在形式上，科学数据在整个生命周期中其数据形式呈动态变化的状态；在时效性上，科学数据具有重复利用的价值，不仅是数据创建者形成研究成果的重要数据支撑，同时还可以为后续的研究者提供参考；在内容上，科学数据安全不仅包含数据本身的安全，还包括相关利益主体的安全 [40]。鉴于科学数据安全管理的重要性和复杂性以及上文对信息安全行为和科学数据安全管理的介绍，笔者把科学数据安全行为定义为：科研工作者在从事科学研究的过程中，为了保证科学数据的真实性、合法性、有效性、机密性、完整性而采取的相关有效措施的行为集合。由于科学研究越来越倾向于以团队的形式开展，科学数据安全行为不仅包含科研人员的个人行为，还包括团队行为。具体来说，科学数据安全行为是科研人员在开展科研活动时，在已有的安全知识储备基础之上，主动采取有利于科学数据安全、规避潜在风险的行为。

2 科学数据安全行为研究现状

科学数据安全行为是信息安全行为的重要分支，通过类比信息安全行为的研究结论，即

知识管理论坛，2019(4):218-231
DOI：10.13266/j.issn.2095-5472.2019.024

认为人为因素是导致信息安全问题的根本原因。在科学数据生命周期中，科学数据在共享发布之前，主要由其拥有者进行管理，在此期间，科学数据安全主要受到个体行为的影响。因此，如何规范科研人员的科学数据安全行为，促进科学数据安全政策的制定，让科研人员在科研过程中对科学数据的安全管理有章可循，对保证科学数据安全，建立科学数据安全管理体系，促进科学数据共享具有重要意义。

国内外关于科学数据安全行为的探讨，嵌入在对数据管理员（data librarian）的角色定位问题中。随着科研模式的变革，图书馆界开始探索数据管理员在科学数据管理过程中应该扮演什么样的角色。数据管理员的职责主要围绕数据管理问题展开，以用户为中心搜集信息资源，保证在多用户环境下信息资源的安全性和可访问性[41]。国外学者如 A. Cremer、R. H. Khan、R. E. Martin 等对数据馆员所需要的职业素养进行了归纳，包括数据的保存与评估、数据监管、开发元数据标准、数据分析与服务、数据安全等[42-44]。国内主要以介绍国外数据管理员的职责为主[45]。在科研活动中，从科学数据的创建到最终利用共享，整个过程都需要科研人员的参与，其中保证科学数据的安全是进行数据分析、利用、共享的基础。虽然相关研究还处于起步阶段，但以数据管理员的职业素养为切入点探讨科学数据管理，为科学数据安全行为的研究提供了借鉴。

从国外高校（如哈佛大学[46]、普林斯顿大学[47]、斯坦福大学[48]等）制定的科学数据管理政策可以看出，科学数据安全问题已经引起广泛关注。虽然不同学校对科学数据安全管理的规定有所差异，但主要集中在以下 3 个方面：①要求研究人员遵守相关协议中的规定；②要求研究人员使用最佳的存储设备和技术，在不给研究人员带来过多负担的情况下安全地保护科学数据；③保护研究对象免受意外披露或不当使用机密数据可能造成的伤害。这些政策从宏观上指明了科学数据安全管理的方向，但为了让科学数据安全管理政策落到实处，则需要科研人员的配合。

在科学数据管理与共享的实践中也会涉及科学数据安全管理问题，如司莉等在研究科学数据管理与共享的需求时对科研人员科学数据管理方式、保存方式等进行了分析，并针对科研是否发生过重要科学数据丢失 / 损毁的现象和重要科学数据丢失 / 损毁的原因等问题进行研究，研究结果显示有 70% 的科研人员发生过重要科学数据丢失 / 损毁的现象，造成重要科学数据丢失 / 损毁的原因主要是误操作或误删除[49]。这也说明科研人员的科学数据安全意识还有待加强，规范科研人员行为以保证科学数据安全至关重要。

3 基于科研过程的科学数据安全行为分析

目前，集科学数据管理、共享、服务于一体的科学数据管理平台主要采用生命周期模型对科学数据进行管理[50]。尹春晓在调研国外高校科学数据管理服务之后总结出，无论是社会科学还是自然科学，其数据管理生命周期都可归纳为以下几个阶段：制定数据管理计划、数据收集与选择、数据描述与组织、数据分析与利用、数据保存与发布、数据发现与获取[51]。可以看出，在科学数据生命周期的不同阶段科研人员的主要工作内容存在差异，这与科研过程具有阶段性息息相关。杨传汶等通过访谈总结了科研过程包括以下 7 个步骤：制定计划、进行试验、收集数据、解释与分析、得出结论、表达与交流、评价与改进[52]。鉴于在科研活动中，科研人员作为主要行为主体，其科研过程的阶段性与科学数据生命周期的划分具有紧密联系。笔者结合科学数据管理政策及相关科学数据管理模型总结了科研过程的一般过程。如图 1 所示。

根据图 1 可以看出，科学研究的一般过程并不像科学数据生命周期模型呈线性，其中，数据存储、数据分析、数据描述处于并行状态。



知识管理论坛
2019 年 第 4 期(总第 22 期)

科学数据在科研过程中一直处于动态变换的状态，这种动态变换让科学数据实现了从数据到知识的跃迁，即原始的科学数据属于数据，处于分析过程中的科学数据属于信息，作为结果的科学数据经过科研人员的解释和说明具有知识的属性。

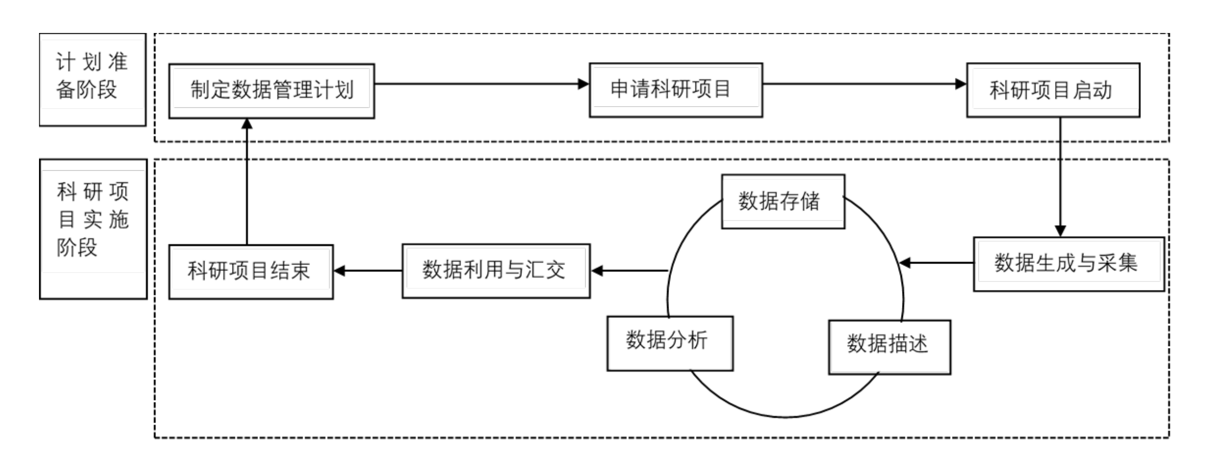


图 1 科学研究的一般过程

由于在科研过程中，科研人员的行为对科学数据安全产生直接影响，笔者将从制定数据管理计划、数据生成与采集、数据存储、数据描述、数据分析、数据利用与汇交几个阶段进行详细阐述，分析在不同阶段可能存在的安全风险，以明确从个人行为角度进行科学数据管理的重点，并提出如何从行为学视角出发加以防范。

3.1 制定数据管理计划

2011 年，美国国家科学基金会（NSF）要求所有基金申请者必须提交对所申请的项目开展的数据管理计划（data management plans，DMP）[53]。随后相关科学数据管理政策也对此做了明确的要求，制定科学数据管理计划已经成了开展科研活动必不可少的工作。科学数据管理计划是科研人员用来描述研究项目中期待获取或生成的数据，对这些数据如何管理、描述、分析、存储，项目结束后采用什么机制来共享和保存数据的书面文件 [54]。完备的数据管理计划是科研人员获取科研资助单位支持的前提，也是其有效管理科研项目所生产科学数据，规避或应对可能产生的各类问题的重要保障 [55]。也就是说，科学数据管理计划有利于科研人员对所从事的科研项目有全局意识，使其更有可能规避科研过程中产生的各类问题。爱丁堡大学科学数据管理政策规定：科学数据管理计划要明确数据采集、管理、完整性、机密性、保留、共享和发布等内容 [56]。目前，科研资助机构主要通过两种方式对科学数据管理计划的撰写提供指导：①提供政策指导，如斯坦福大学提出的数据管理计划指南 [50]；②提供技术服务，如英国数据监管中心（DCC）提供了支持标准化的工具 DMPOnline[57]，美国加州福尼亚数字图书馆提供支持和服务的数据管理计划制定工具 DMPTool[58]。笔者通过调研国内外相关科学数据管理计划的政策文献 [59-61]，归纳了科学数据管理计划包含的主要内容：①对科研项目进行描述；②数据的组织和存储；③数据的获取、分享和重用说明；④数据存档说明。其每个部分包含的具体内容如表 1 所示。

3.2 科学数据生成与采集

科学数据生成与采集过程是建立在系统的数据收集方法之上，针对设计的研究变量收集数据，为证明研究问题、验证假设和评估结果服务。虽然数据收集方法因学科差异而有所不同，但对数据质量都同样重视。由科学数据安

知识管理论坛，2019(4):218–231
DOI：10.13266/j.issn.2095–5472.2019.024

全行为的定义可知，保证科学数据的真实性、合法性是数据收集过程中的核心问题。收集不合适的数据可能带来以下影响：无法准确回答研究问题；无法进行重复验证；造成资源浪费；对其他研究人员形成误导；对人类研究者或动物主体造成伤害。除此之外，如果将由这种不合适的数据用于建立公共政策的数据支撑，则可能造成严重的不良影响。为了保证数据的真实性、合法性和完整性，G. L. Knatterud 提出可以从质量保障（quality assurance）和质量控制（quality control）两方面入手[62]。质量保证是指在数据收集之前进行的活动，可以通过以下方法对科学数据的收集提供保障：①编写数据收集程序手册；②提前了解收集数据需要了解的领域知识；③掌握该领域现有的政策和标准；④学习数据收集工具的使用方法。质量控制是指在数据收集期间和数据收集之后进行的活动，在这个过程中，需要数据收集者明确其职责，严格按照数据收集步骤进行，并建立反馈机制对其进行评估，发现存在的问题后，及时进行修正调整。针对质性研究，为了验证数据质量，可以采用三角互证的研究方法[63]，利用社会期望量表来减小应答偏差对实验数据效度的影响[64]。

表 1　数据管理计划应包含的内容

描述项目	内容
科研项目	研究目的；数据的来源；如何获取数据；数据的类型；数据管理的权责划分
数据的组织和存储	元数据的创建、管理、存储说明；文件命名规则说明；本地存储和备份过程说明；数据读取说明（查看数据需要用到什么工具和软件）
数据的获取、分享和重用	谁有权访问这些数据；可以通过哪些途径获取；数据共享说明（是否涉及隐私、道德或保密问题）；这项研究的知识产权说明；是否允许数据重用；数据的授权（是否允许商业用途）
数据存档	数据长期保存说明（是否会存储在数据存储中心，如果不是，将如何存储数据）；数据的保存时间

3.3 存储数据

科学数据的存储与维护是实现科学数据共享利用、发挥其最大价值的重要基础[65]。存储数据之所以重要，是因为许多科学数据是独一无二的，一旦被毁坏将可能永远丢失[66]。在我国，由于项目资助机构对科学数据提交没有强制要求，导致大部分数据分散在课题组或个人手中，使科学数据面临损坏或永远丢失的危险[67]。对于数据的存储，笔者通过调研国内外科学数据保存政策，总结出科学数据存储时应该考虑以下几个方面：

（1）数据的命名规则。文件名是文件的主体标识符，规范的文件名可以为文件的内容、状态和版本提供有用的提示，有助于对文献进行分类和排序。南安普顿大学和英国数据服务中心制定了数据文件命名规则可供参考[68-69]：①文件名应该做到简洁且意义明确；②整个科研项目的文件应该按照同一个命名规则，保证文件的一致性；③避免使用特殊符号（如 & ? ！）；④避免文件名过长；⑤借助连字符“-”或下划线“_”分隔文件名中的元素，使其具有更好的可读性；⑥为针对特定应用程序的代码文件保留文件扩展名（如 .doc、.xls、.tif 等）；⑦在文件名中包含文件的版本控制。

（2）数据的存储位置。选择合适的存储位置对科学数据管理和利用带来很大的便利。通常可以用来存储科学数据的有以下 5 类[70]：①机构存储，一般指学校或科研机构的数据存储中心；②本地存储，如个人电脑；③云存储，如 DropBox、Google Drive 等；④便携式存储设备，如 USB 存储器、DVD 等；⑤纸质存储。每一种存储方式都有优缺点，科研人员在选择存储时，还应该参照项目中产生科学数据的大小、数据的保密级别、访问需求（是否提供远程访问）等因素选择合适的存储方式选择合适的数据存储位置。

（3）数据文件存储的格式。恰当的文件存储格式可以避免数据失效并便于数据重用[71]。在选择文件格式时，必须考虑：数据的分析方法、相关标准、软件和硬件的兼容性、是专属软件还是开源软件等问题[72]。在数据收集和分析过程中，研究人员可能会选择比较方便的通用格式，但为了使数据能长期存储，便于未来进行数据访问，建议将数据存储为开放格式。斯坦福大学图书馆建议的数据存储格式[73]如表2所示。

（4）备份策略。备份数据可以确保原始文件因一些意外原因（如硬件故障、计算机病毒感染等）而丢失时，可以从备份的副本中恢复原始数据文件，从而减少损失。选择备份策略时要考虑本地存储情况、数据的价值以及环境的风险级别。英国数据服务中心提出的在制定数据备份策略时应该考虑的问题及建议[74]如表3所示。

表 2 科学数据存储格式

文件类型	建议的存储格式
文本文件	XML、PDF/A、HTML、ASCII、UTF-8
音频文件	WAVA、AIFF、MP3、MXF
压缩文件	TAR、GZIP、ZIP
数据库文件	XML、CSV
地理空间数据	SHP、DBF、GeoTIFF、NetCDF
运动图像	MOV、MPEG、AVI、MXF
统计数据	ASCII、DTA、POR、SAS、SAV
静态图像	TIFF、JPEG2000、PDF、PNG、GIF、BMP
表格数据	CSV
网页	WARC

表 3 制定数据备份策略

备份项目	具体内容
备份的内容	根据机构是否可以恢复系统的具体情况，考虑是备份特定的数据还是备份整个系统。如果可以恢复系统，则只需要备份数据文件即可
是否包含敏感信息	如果数据文件涉及个人敏感信息，应该只创建所需的最少的备份副本，并对数据进行加密
备份的频次	建议在每次更改数据后进行备份，或定期备份，或者借助工具对关键数据文件进行自动备份
可实现自动备份的工具	微软的synctoy、苹果Mac的Time Machine等
备份存储介质的选择	少量的数据文件可以使用CD/DVD等，如果是大量的数据文件，则建议使用移动硬盘驱动器等
备份数据的保存格式	采用适合长期保存的文件格式，可参照表2
增量备份还是差异备份	增量备份是指备份自上一次备份之后有变化的数据，建议使用可移动设备进行备份。差异备份是指备份上一次完全备份之后有差异的数据，建议使用固定介质（如硬盘驱动器）进行备份
如何组织备份数据	通过建立标签体系，定期验证备份的文件，通过与原始文件进行比较，检查其完整性。
备份数据的存储位置	一般数据可以存储在网络云盘上，对于涉及个人信息的数据建议脱机存储

（5）数据存储的安全性。科学数据的安全不仅包括科学数据本身的安全，还包括相关利益主体的安全。为了确保数据本身的存储安全性，可以采取以下措施：①控制访问权限。可以通过对数据进行加密来控制访问权限。除了对数据进行加密之外，同时还要对存储有科学数据的存储设备（如笔记本电脑、台式机和移动设备）进行加密，降低未经授权的访问导致存储科学数据的设备暴露的风险[75]。②对数据进行分级。对于涉及人类参与者的科学数据，可以通过对数据进行分级。哈佛大学科学数据管理中心[76]提出将科学数据分成5个等

知识管理论坛，2019(4):218–231
DOI：10.13266/j.issn.2095–5472.2019.024

级，并对不同级别的数据提出了存储方案，如表 4 所示：

表 4 哈佛大学的数据分级制度

级别	描述
Level5：极其敏感的数据	如果泄露会对个人或学校造成严重伤害的数据，如可识别个人身份的医疗信息、遗传信息
Level4：非常敏感的数据	如果泄露可能会对个人或学校造成严重伤害的数据，如可识别的财务信息、国家安全信息、可获取机密信息的密码
Level3：敏感或机密额数据	如果泄露可能会对个人或学校造成重大风险的数据，包括机构人事记录，财务记录，个人捐助信息，除第四、五等级之外受法律保护的个人信息
Level2：保密的良性数据	泄露不会造成重大损害，学校还是会选择保密的数据，包括未发表的科研成果以及第三、四等级之外的知识产权、论文初稿、专利申请文件
Level1：非保密的数据	公共信息，包括没有身份标识的科研数据、已发表的研究数据、学校公开信息等

3.4 数据描述

司莉通过实证研究发现，有 50% 以上的科研人员表示科学数据管理中的问题有不知道如何为数据创建索引以及不记得数据的存储位置 [77]。对科学数据进行描述可以使科研人员更容易对数据进行定位，方便对数据的理解和重用 [78]。数据描述是指利用标准化、规范化的术语对数据进行描述 [79]。通常使用元数据对科学数据进行描述。M. Wilkinson 等从 4 个维度提出了科学数据元数据的准则，即易发现、易获取、可操作、可重用 [80]。从具体应用层面来看，国外一些机构和大学图书馆制定了元数据的标准 [81]，如表 5 所示：

表 5 数据描述的基本要素

描述项目	描述
题名	数据集或科研项目的名字
作者	创建数据的机构或研究人员的名字
标识符	用于标识数据的编号
主题	描述数据的关键词或短语
资助者	资助研究项目的组织或机构
权限	其他研究者如何获取数据的说明
语言	研究成果所包含数据资源的语言
日期	与数据有关的关键信息，包括项目开始日期、结束日期、发布日期、所经历的时间周期。其他的日期如数据更新日期、维护周期等
地址	记录与数据有关的物理空间信息
方法	数据的收集方法，包括所使用的设备和软件、实验协议等
数据处理	关于数据处理的过程和步骤
资源	引用的其它数据资源，记录其数据来源和获取方法
文件列表	与该科研项目有关的所有数据文件名单
文件格式	对数据格式的说明
文件结构	数据文件的层级
变量列表	数据文件中所涉及的所有变量列表
代码清单	对文件名或数据文件中的变量使用的代码或缩写的解释
版本	每个文件发布的时间
校验	参照备份的文件，测试文件是否随着时间的推移而更改

针对不同的学科，也有相应的元数据标准 [82-83] 如表 6 所示：

表 6 不同学科的元数据标准

学科	元数据标准
社会科学数据	社会科学标准（DDI）
地理空间数据	数字地理元数据内容标准（CSDGM） 地理信息国际标准（ISO19115:2003）
人文数据	都柏林核心集（DC）
多媒体数据	数字图像的元数据标准（NISO Z39.87-2002）
天文数据	天文可视化元数据标准（AVM）

3.5 数据分析

数据分析是科研人员利用科学的数据分析方法把原始的科学数据转化为科研成果的阶段，是科研人员对科学数据操作最频繁的阶段，最容易出现的科学数据安全风险包括误操作或误删除、存储设备故障等。由图 1 可以看出，在科研活动中，数据存储、数据分析、数据组织三者是相互交织的过程，具体来说数据存储和数据组织是为了数据分析做准备，数据分析和数据组织之后必须要对数据进行存储，数据组织是为了更好地进行数据存储和数据分析。在进行数据分析时，科学数据的形式处于动态变化的状态，为了保证数据分析的顺利进行，应该根据数据描述和数据存储规则对数据进行描述和存储。

3.6 数据利用与汇交

关于科学数据共享的实证研究数据显示，42.82% 的科研人员表示在科研项目结束后，科学数据分散在项目组成员手中 [84]。这表明一部分科学数据并没有被利用起来，科研人员汇交科学数据的积极性还有待提高。对科学数据进行汇交是科学数据从个人管理层面进入数据流通系统的重要环节，按照相关标准对科学数据进行整理，提交到数据共享中心。数据汇交是科研人员对自己在科研过程中产生的科学数据的集中整理，确保科学数据的有效性和完整性。根据《科学数据管理办法》的第十三条规定，由政府预算资金资助的各级科技计划项目所形成的科学数据，应由项目牵头单位汇交到相关科学数据中心 [9]。从科研人员的角度来看，要考虑的问题是把数据交到什么地方去？以及交什么？如何交？针对这一系列问题，国外科研项目管理机构、学术期刊及数据机构制定了科学数据汇交政策 [85]。我国于 2008 年 3 月首次由科技部颁布了《国家重点基础研究发展计划资源环境领域项目数据汇交暂行办法》[86]。要求把科学数据汇交到相应的数据汇交中心。鉴于科学数据的复杂性，不同学科领域建立了科学数据汇交管理办法，部分领域的科学数据汇交管理办法见表 4，同时，部分领域如气象、测绘、水文的数据共享管理办法也包含数据汇交细则。这些管理办法对数据汇交的组织管理、汇交内容、数据汇交计划、数据汇交流程、数据管理、权益保护、监督与信用管理等都做了规定。数据汇交的内容包括科研项目生成或采集的原始数据以及对原始数据进行处理和加工后形成的数据。为了保证科学数据汇交工作的顺利进行，科学数据管理中心引入流程管理思想来规范数据汇交与管理 [87]。及时汇交和共享科学数据资源，既是国家科技投入的直接效益体现，也是促进这些数据更好地被归档、存储、共享和开发利用的重要途径 [79]。所以科研人员在科研项目结束后，应该充分了解相关数据汇交管理办法，并以此为基础对科学数据进行整理和筛选，形成统一规范的数据格式，及时地提交到数据共享中心。

表 7 部分科学数据共享中心及其数据汇交管理办法

科学数据共享中心	科学数据汇交管理办法
国家人口与健康科学数据共享中心	《中医药科研课题数据汇交管理办法》[88]
国家农业科学数据共享中心	《农业科学数据汇交管理办法》[89]
地震科学数据共享中心	《人工地震科学数据汇交与共享管理办法（试行）》[90]
资源环境数据云平台	《国家重点基础研究发展计划资源环境领域项目数据汇交暂行办法》[86]
南海海洋数据中心	《中国科学院南海海洋研究所数据汇交管理暂行办法》[91]
国家材料科学数据共享中心	《材料科学数据提交格式规范》[92]

知识管理论坛，2019(4):218−231
DOI：10.13266/j.issn.2095−5472.2019.024

4 结语

科学数据安全不仅涉及科学数据本身的安全，也包括相关利益主体（个人、组织、国家）的安全。在科学数据开放共享的时代，科学数据安全管理是实现科学数据共享的基础。虽然相关政策对科学数据管理平台、组织机构（如个高校图书馆、科研资助机构）以及科研工作者划定了他们的职责和义务。但总体而言，国家法律的条例是概括性的、原则性的。如何将其分别具体细化到各类行为主体并实施于科学数据安全管理实践，则亟需科学数据安全行为的及时跟进。根据科学研究的一般过程，科学数据在共享之前，其安全主要受到科研人员行为的影响，所以笔者探讨的科学数据安全行为主要针对科研人员。根据上述分析，为保障科学数据安全，在科研过程的不同阶段，由于科研工作者的工作内容的差异导致其所采取的措施也不同。在制定科学数据管理计划阶段，需要研究者对科研项目所产生的数据有大致了解，并对如何管理科学数据做出规划，才能在后续阶段对科学数据安全管理有一个宏观的把控；在数据收集与采集阶段，为保障科学数据质量，研究者需要在数据收集前做相应的准备工作，在数据收集过程中采取相关措施对数据质量进行控制；在数据存储阶段，可以从数据的命名规则、存储位置、数据文件存储格式、数据备份策略等几个方面保障科学数据安全；在数据组织阶段，可以参照相关学科元数据标准对数据进行描述；在数据分析阶段，则应该选择科学的数据分析方法对数据进行分析，同时参照数据存储和数据组织的相关准则对数据进行操作；在数据利用与汇交阶段，需要对数据进行整理筛选，按照相关科学数据汇交管理办法汇交到科学数据管理中心，实现科学数据的共享。

科学数据安全行为是科研人员在科研活动中为保证科学数据及相关利益主体安全而采取的主动防御措施，是基于个人行为从微观层面对科学数据提供的保障。规范科研人员的安全行为是构建科学数据安全体系必不可少的步骤。笔者根据现实需求，整合信息安全行为、科学数据安全管理的相关理论和政策，基于科研过程不同阶段的工作内容的差异分析了可能存在的安全风险及应该如何加以防范，以期为科研人员在科研过程中实现对科学数据安全管理提供指导和建议，同时为科研管理机构制定科学数据安全管理计划提供参考。本文不足之处在于，现阶段关于科学数据安全的研究成果主要针对科学数据管理中心，这些政策及建议对科学数据安全行为的适用性还有待实证检验，这也是下一步需要继续研究的方向。

参考文献：

[1] 中国科学院．第四届（2017）科学数据大会会议通知 .[EB/OL]. [2018-10-20]. http://www.cas.cn/xs/201707/t20170711_4608195.shtml.

[2] 中华人民共和国中央人民政府．科学数据是重要战略资源 [EB/OL]. [2018-11-24]. http://www.gov.cn/zhengce/2018-04/06/content_5280211.htm.

[3] 黄如花，王斌，周志峰．促进我国科学数据共享的对策 [J]. 图书馆，2014(3): 7-13.

[4] World Data System. Data sharing principles[EB/OL]. [2018-10-08]. https://www.icsu-wds.org/.

[5] Australian National Data Service. Research data Australian[EB/OL]. [2018-10-08]. https://www.ands.org.au/.

[6] Digital Curation Centre[EB/OL]. [2018-10-08]. http://www.dcc.ac.uk/.

[7] The University of Edinburgh. Research resources[EB/OL]. [2018-10-08]. https://www.ed.ac.uk/information-services/library-museum-gallery/crc/research-resources.

[8] Princeton University. Research data security[EB/OL]. [2018-10-10]. https://ria.princeton.edu/research-data-security.

[9] 中华人民共和国人民政府．国务院办公厅关于印发科学数据管理办法的通知 [EB/OL]. [2018-08-12]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.

[10] The Stationery Office. Data Protection Act 1998[EB/OL]. [2018-10-10]. http://www.legislation.gov.uk/ukpga/1998/29/pdfs/ukpga_19980029_en.pdf.

[11] BBSRC.BBSRC data sharing policy[EB/OL]. [2018-04-18]. http://www.bbsrc.ac.uk/documents/data-sharing-policy-pdf/.

[12] Health Information Privacy. Summary of the HIPAA se-

curity rule[EB/OL]. [2018-10-20]. https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html.

[13] Research Ethics and Compliance. Data security guidelines[EB/OL]. [2018-10-20]. https://research-compliance.umich.edu/data-security-guidelines.

[14] IDG—Insights Intent and Engagement.2018 Global state of information security survey[EB/OL]. [2018-12-02]. https://www.idg.com/tools-for-marketers/2018-global-state-information-security-survey/.

[15] Joint Information Systems Committee. Security of research data[EB/OL]. [2018-10-10]. https://www.jisc.ac.uk/guides/data-protection-and-research-data/security-of-research-data.

[16] Euronews. The Facebook data leak: what happened and what's next[EB/OL]. [2018-10-12]. https://www.euronews.com/2018/04/09/the-facebook-data-leak-what-happened-and-what-s-next.

[17] 中华人民共和国中央人民政府 . 科学数据，如何科学 管 理 [EB/OL]. [2018-08-29]. http://www.gov.cn/zhengce/2018-04/08/content_5280429.htm.

[18] 刘桂锋 , 魏悦 , 钱锦琳 . 高校科研数据管理与共享政策的案例与执行模型研究 [J]. 图书馆论坛 , 2018, 38(11): 27-34.

[19] 曾忠平 , 杨哲 , 刘春梅 . 用户信息安全行为研究述评 [J]. 情报杂志 , 2014, 33(12): 184-188.

[20] 李晶 . 信息安全行为研究现状与发展动态述评 [J]. 图书情报工作 , 2014, 58(24): 126-130.

[21] WOOD C C, BANKS JR W W. Human error: an overlooked but significant information security problem[J]. Computers and security, 1993, 12 (1): 51-60.

[22] MCCAULEY-BELL P R, CRUMPTON L L. The human factors issues in information security: what are they and do they matter[C]//Proceedings of the human factors and ergonomics society annual meeting. Los angeles: Sage publications, 1998:439-443.

[23] ANDERSON C L, AGARWAL R. Practicing safe computing: a multimedia empirical examination of home computer user security behavioral intentions[J]. Management information systems quarterly, 2010, 34 (3): 613-643.

[24] SCHULTZ E. The human factor in security[J]. Computers and security, 2005, 24 (6): 425-426.

[25] STANTON J M, CALDERA C, GUZMAN I, et al. Behavioral information security: an overview, research agenda, and preliminary results[C]//The security conference. Las vegas: Nevada, 2003:23-24.

[26] DZAZALI S, SULAIMAN A, ZOLAIT A H. Information security landscape and maturity level: case study of Malaysian Public Service (MPS) organizations[J]. Government information quarterly, 2009, 26 (4): 584-593.

[27] LIANG H, XUE Y. Understanding security behaviors in personal computer usage: a threat avoidance perspective[J]. Journal of the Association for Information Systems, 2010, 11 (7): 394-413.

[28] WIRTZ J, LWIN M O, WILLIAMS J D. Causes and consequences of consumer online privacy concern[J]. International journal of service industry management, 2007, 18 (4): 326-348.

[29] 黄鑫 , 邓仲华 . 数据密集型科学研究的需求分析与保障 [J]. 情报理论与实践 , 2017, 40(2): 66-70.

[30] 赵华 , 朱亮 , 鲜国建 , 等 . 数据监护现状分析及对我国农业科学数据监护的启示 [J]. 数字图书馆论坛 , 2017(11): 9-14.

[31] 沈怡 . 科研数据实践的实证研究对数据管理和共享的重要意义 : 个案回顾、反思与前瞻 [J]. 图书情报知识 , 2018(4): 102-108.

[32] LEE D J, STVILIA B. Practices of research data curation in institutional repositories: a qualitative view from repository staff[J]. PLOS ONE, 2017, 12 (3): 1-44.

[33] The Digital Curation Centre. What is digital curation?[EB/OL]. [2018-11-12]. http://www.dcc.ac.uk/digital-curation/what-digital-curation.

[34] 吴金红 , 陈勇跃 . 面向科研第四范式的科学数据监管体系研究 [J]. 图书情报工作 , 2015, 59 (16): 11-17.

[35] 宋筱璇 , 王延飞 , 钟灿涛 . 国内外科研数据安全管理政策比较研究 [J]. 情报理论与实践 , 2016, 39 (11): 10-16.

[36] 丁培 . 国外大学科研数据管理政策研究 [J]. 图书馆论坛 , 2014, 34 (5): 99-106.

[37] 张瑶 , 顾立平 , 杨云秀 , 等 . 国外科研资助机构数据政策的调研与分析 —— 以英美研究理事会为例 [J]. 图书情报工作 , 2015, 59 (6): 53-60.

[38] 郝世博 , 徐文哲 , 唐正韵 . 科学数据共享区块链模型及实现机理研究 [J]. 情报理论与实践 , 2018(11): 57-62.

[39] 司莉 , 邢文明 . 国外科学数据管理与共享政策调查及对我国的启示 [J]. 情报资料工作 , 2013(1): 61-66.

[40] 马海群 , 蒲攀 . 国内外开放数据政策研究现状分析及我国研究动向研判 [J]. 中国图书馆学报 , 2015, 41(5): 76-86.

[41] LISCOUSKI J. The Data Librarian: introducing the Data Librarian[J]. Journal of analytical methods in chemistry, 1900, 19(6): 199-204.

知识管理论坛，2019(4):218–231
DOI：10.13266/j.issn.2095–5472.2019.024

[42] CREAMER A, MORALES M E, CRESPO J, et al. An assessment of needed competencies to promote the data curation and management librarianship of health sciences and science and technology librarians in New England[J]. Journal of eScience librarianship, 2012, 1 (1): 4.

[43] KHAN H R, DU Y. What is a data librarian?: A content analysis of job advertisements for data librarians in the United States Academic Libraries[J]. International federation of library associations and institutions, 2018(7): 1-9.

[44] MARTIN E R. Highlighting the informationist as a data librarian embedded in a research team[J]. Journal of eScience librarianship, 2013, 2 (1): 1-3.

[45] 胡绍君 . 面向科研数据管理的高校学科馆员能力建设研究 [J]. 图书情报工作 , 2016, 60(22): 74-81.

[46] Harvard University. Harvard research data security policy[EB/OL]. [2019-05-16]. https://vpr.harvard.edu/pages/harvard-research-data-security-policy.

[47] Princeton University. Research data security[EB/OL]. [2019-05-16]. https://ria.princeton.edu/research-data-security.

[48] Stanford Medicine. Research and security[EB/OL]. [2019-05-16]. https://med.stanford.edu/irt/security/research.html.

[49] 司莉 , 邢文明 . 科学数据管理与共享的理论与实践 [M]. 武汉 : 武汉大学出版社 , 2017: 81.

[50] 杨林 , 钱庆 , 吴思竹 . 科学数据管理生命周期模型比较 [J]. 中华医学图书情报杂志 , 2016, 25 (11): 1-6.

[51] 尹春晓 . 高校科学数据管理嵌入式服务模式探索 [J]. 情报资料工作 , 2017, 38 (2): 77-82.

[52] 杨传汶 , 徐坤 . 基于生命周期的动态科学数据服务模式研究 [J]. 图书馆论坛 , 2015, 35 (10): 82-87.

[53] Nation Science Foundation. NSF data management plan requirements[EB/OL]. [2018-10-23]. https://www.nsf.gov/bfa/dias/policy/dmp.jsp.

[54] Stanford Libraries. Data management plans[EB/OL]. [2018-09-01]. https://library.stanford.edu/research/data-management-services/data-management-plans.

[55] 黄国彬 , 邸弘阳 , 张莎莎 , 等 . 数据管理计划工具 DMPTool 的服务体系剖析 [J]. 图书情报工作 , 2018, 62(4): 37-43.

[56] 王海彪 , 卫军朝 . 科学数据管理关键因素研究 —— 基于爱丁堡大学科学数据管理实践及启示 [J]. 图书馆杂志 , 2017, 36(1): 20-26.

[57] Digital Curation Centre.DMPonline[EB/OL]. [2018-08-13]. https://dmponline.dcc.ac.uk/.

[58] University of California Curation Center. DMPTool[EB/OL]. [2018-11-10]. https://dmptool.org/.

[59] MIT Libraries. Write a data management plan[EB/OL]. [2018-11-12]. https://libraries.mit.edu/data-management/plan/write/.

[60] Stanford Libraries. About data management plans (DMPs)[EB/OL]. [2018-11-12]. https://library.stanford.edu/research/data-management-services/data-management-plans.

[61] Medical Research Council. What is a data management plan?[EB/OL]. [2018-11-12]. https://mrc.ukri.org/documents/pdf/what-is-a-data-management-plan/.

[62] KNATTERUD G L, ROCKHOLD F W, GEORGE S L, et al. Guidelines for quality assurance in multicenter trials: a position paper[J]. Controlled clinical trials, 1998, 19 (5): 477-493.

[63] DENSCOMBE M. 怎样做好一项研究 : 小规模社会研究指南 [M]. 陶保平 , 译 . 上海 : 上海教育出版社 , 2011:115-119.

[64] Northern Llinois University. Data collection[EB/OL]. [2018-12-11]. https://ori.hhs.gov/education/products/n_illinois_u/datamanagement/dctopic.html.

[65] 司莉 , 封洁 . 科学数据的保存与维护 : 国际组织的动向 [J]. 图书馆 , 2015(4): 6-10.

[66] Joint Information Systems Committee. Why preserve digital data?[EB/OL]. [2018-08-21]. http://www.dcc.ac.uk/digital-curation/why-preserve-digital-data.

[67] 谢春枝 , 燕今伟 . 国内外高校科学数据管理和机制建设研究 [J]. 图书情报工作 , 2013, 57 (6): 12-17.

[68] Southampton Library. Research data management: file naming[EB/OL]. [2018-12-01]. http://library.soton.ac.uk/researchdata/filenaming.

[69] UK Data Service. File names[EB/OL]. [2018-12-01]. https://www.ukdataservice.ac.uk/manage-data/format/organising.

[70] University of Liverpool. Storing your research data[EB/OL]. [2018-12-01]. https://www.liverpool.ac.uk/library/research-data-management/storing-your-research-data/.

[71] University of Concordia. Data storage and file formats[EB/OL]. [2018-12-03]. https://library.concordia.ca/help/data/data-storage.php.

[72] The University Of Western Australia Research data preservation formats[EB/OL]. [2018-12-03]. http://guides.library.uwa.edu.au/c.php?g=325196&p=2178568.

[73] Stanford Libraries. Best practices for file formats[EB/OL]. [2018-12-04]. https://library.stanford.edu/research/data-management-services/data-best-practices/best-practices-file-formats.

[74] UK Data Service. Data backup[EB/OL]. [2018-12-04]. https://www.ukdataservice.ac.uk/manage-data/store/backup.aspx.

[75] Fordham University. Research data security guidelines[EB/OL]. [2018-12-04]. https://www.fordham.edu/info/24327/guidelines_and_procedures/9959/research_data_security_guidelines/1.

[76] Harvard University. Harvard research data security policy[EB/OL]. [2018-12-05]. https://vpr.harvard.edu/pages/harvard-research-data-security-policy.

[77] 司莉 , 邢文明 . 科学数据管理与共享的理论与实践 [M]. 武汉 : 武汉大学出版社 , 2017: 60.

[78] Southampton Library. Research data management: describing[EB/OL]. [2018-12-05]. http://library.soton.ac.uk/researchdata/description.

[79] 涂志芳 . 科学数据出版生态系统与质量控制体系构建 [J]. 图书与情报 , 2019(1): 125-134.

[80] WILKINSON M, DUMONTIER M, AALBERSBERG I. The FAIR guiding principles for scientific data management and stewardship[J]. Scientific data, 2016(3): 167-172.

[81] Oulun Yliopisto. Research data guide[EB/OL]. [2018-12-04]. http://libguides.oulu.fi/Researchdata/Data_documentation.

[82] The University of Western Australia. Metadata standards[EB/OL]. [2018-12-04]. https://guides.library.uwa.edu.au/c.php?g=325196&p=2178564.

[83] The Digital Curation Centre. List of metadata standards[EB/OL]. [2018-12-04]. http://www.dcc.ac.uk/resources/metadata-standards/list.

[84] 司莉 , 邢文明 . 科学数据管理与共享的理论与实践 [M]. 武汉 : 武汉大学出版社 , 2017: 61.

[85] 王卷乐 , 祝俊祥 , 杨雅萍 , 等 . 国外科技计划项目数据汇交政策及对我国的启示 [J]. 中国科技资源导刊 , 2013(2): 17-23.

[86] 国家重点基础研究发展计划资源环境领域项目数据汇交暂行办法 [EB/OL]. [2018-12-08]. http://www.973geodata.cn/files/1.pdf.

[87] 宋转玲 , 刘海行 , 葛人峰 , 等 . 国家自然科学基金委员会共享航次调查数据汇交体系构建 [J]. 中国科学基金 , 2014, 28(6): 455-459.

[88] 国家科技基础条件平台 . 中医药科研课题数据汇交管理办法 [EB/OL]. [2018-12-10]. http://dbcenter.cintcm.com/content.jsp?id=189.

[89] 国家农业科学数据共享中心 . 农业科学数据汇交管理办法 [EB/OL]. [2018-12-10]. http://fishery.agridata.cn/ch_law.asp.

[90] 陕西省地政局 . 人工地震科学数据汇交与共享管理办法（试行）[EB/OL]. [2018-12-10]. http://www.shxdzj.gov.cn/n1172/n1175/c532930/part/532931.pdf.

[91] 中国科学院南海海洋研究所 . 中国科学院南海海洋研究所数据汇交暂行办法 [EB/OL]. [2018-12-08]. http://odc.scsio.ac.cn/images/123.pdf.

[92] 国家材料科学数据共享网 . 材料科学数据提交格式规范 [EB/OL]. [2018-12-10]. http://matsec.ustb.edu.cn/uploadFiles/shujutijiao.pdf.

作者贡献说明：

杨　燕：确定论文框架，进行资料收集和分析，撰写论文；

阮建海：指导论文写作修改论文。

知识管理论坛，2019(4):218-231
DOI：10.13266/j.issn.2095-5472.2019.024

Research on Scientific Data Security Behavior Based on Scientific Research Process

Yang Yan[1] Ruan Jianhai[2]
[1] School of Computer and Information Science, Southwest University, Chongqing 400715
[2] Southwest University Library, Chongqing 400715

Abstract: [Purpose/significance] This paper defines the concept of scientific data security behavior, and combines different stages of research activities to analyze how to ensure the safety of scientific data in the scientific research process. It is of great significance to standardize the scientific data security behavior of scientific researchers, promote scientific data security management, formulate scientific data security policies, and improve the scientific data security management system. **[Method/process]** The network research method was adopted to integrate the definition of information security behavior and scientific data security management, and then, the concept of scientific data security behavior was proposed. On this basis, starting from the theory, this paper combined with the practical experience of scientific data management and sharing to explained that way to ensure the security of scientific data through personal behavior. **[Result/conclusion]** At this stage, scientific data security management is aimed at scientific data management platform. Research on scientific data security from the perspective of behavior is still in its infancy. Based on the scientific research process, this paper analyzes how to standardize the behavior of scientific researchers to improve the security of scientific data from two aspects. First, in the stage which existence of scientific data security risk, combined with the policy of scientific data management, this paper proposes a targeted improvement strategy. Second, for the stage where there is no obvious security risk, suggestions are put forward which based on the practical experience of scientific data management at home and abroad.

Keywords: scientific data security behavior　scientific data　information security behavior　scientific data management